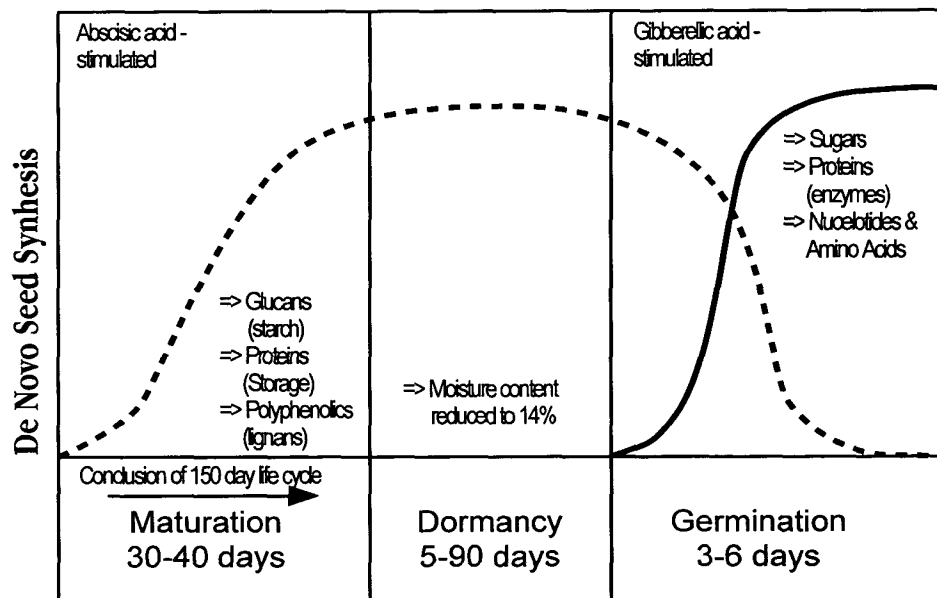
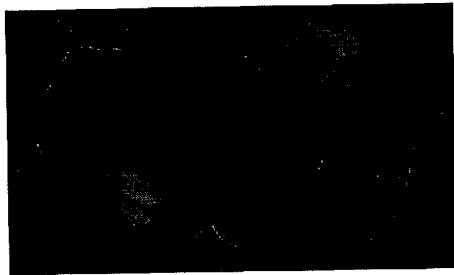


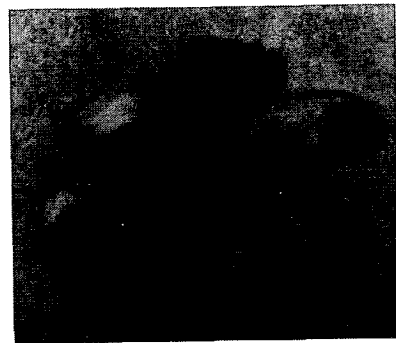
**Fig. 1A**



**Fig. 1B**



**Fig. 1C**



**Fig. 1D**



**Fig. 1E**



**Fig. 1F**

102030 2224300

10 20 30 40 50 60 70  
GGTACCCATC TAATACATTA ATAACAAGAG AGAGAATGGA TAATGCAATT ATTTATTTTT ATGGGAGGCT  
CCATGGGTAG ATTATGTAAT TATTGTTCTC TCTCTTACCT ATTACGTAA TAAATAAAAA TACCCTCCGA

80 90 100 110 120 130 140  
ATATTTTTAT CGGATTTTAG TAAATAACGG GGCAATTCGG TACTTAGGTA AAGCTACGTA TGACTATCGC  
TATAAAAATA GCCTAAAATC ATTTATTGCC CCGTTAAGCC ATGAATCCAT TTCGATGCAT ACTGATAGCG

150 160 170 180 190 200 210  
TACCGCTACG GTAGTTGAAT TGGAATTCCT CGATAGCATC TGTTGTGTTG TTGCAGTTAG GGTACTTGAA  
ATGGCGATGC CATCAACTTA ACCTTAAGAA GCTATCGTAG ACAACACAAC AACGTCAATC CCATGAACCT

220 230 240 250 260 270 280  
TAGCTCCAGC CGTGAAAACG AGGGGTTTTC GCAGGTTTTA TAGGATTGCC AAGTTAGACT AGGGCAATTC  
ATCGAGGTCG GCACTTTTGC TCCCCAAAAG CGTCCAAAAT ATCCTAACGG TTCAATCTGA TCCCCTTAAG

290 300 310 320 330 340 350  
ATGTTACCGG TATTGTGTAG TATATGAAAA AGGAGATCTC CCAAACAATT TATAATTTTG TATAAGGGAG  
TACAAGTGCC ATAACACATC ATATACTTTT TCCTCTAGAG GGTTTGTAA ATATTAAAC ATATTCCCTC

>AT-rich\_region\_  
|

360 370 380 390 400 410 420  
AAATCGAACT TGAGGTGTCT AATTCACCAA CCGAGCTACT CCCTCCGTTT CATATATGTA TATACATATA  
TTTAGCTTGA ACTCCACAGA TTAAGTGGTT GGCTCGATGA GGGAGGCAAA GTATATACAT ATATGTATAT

430 440 450 460 470 480 490  
TACGTATATA TACGTATATA CACATATACG TATATACATA TATGGTATAT ACATATATAT ATATATATAT  
ATGCATATAT ATGCATATAT GTGTATATGC ATATATGTAT ATACCATATA TGTATATATA TATATATATA

500 510 520 530 540 550 560  
ATATATATAT ATGTGTGTGT GTGTATGTGG GGTGGCAATG CTAAAAAGTT TTATAATATG AACGGATGAA  
TATATATATA TACACACACA CACATACACC CCACCGTTAC GATTTTTCAT AATATTATAC TTGCCTACTT

570 580 590 600 610 620 630  
GTACTATCCA CTAAGTCCCT ATAGTTTCTT GGCAGTGTGT AGTATACGAA TGCACAATTA TATCCATAAA  
CATGATAGGT GATTCAGGGA TATCAAAAAG CCGTGACACA TCATATGCTT ACGTGTTAAT ATAGGTATTT

640 650 660 670 680 690 700  
ATTGATATTA TATATTCGTC GCGACGAAAA TAAAGACATA ATATTCGGTA TACCATTTAT CCACGATATA  
TAACATAAT ATATAAGCAG CGCTGCTTTT ATTTCTGTAT TATAAGCCAT ATGGTAAATA GGTGCTATAT

710 720 730 740 750 760 770  
TCTAAATTCC ACTGATATAT CTAAATTCCA CTTGATCCCT TTTATGGATA AATTCTGGAT AACAAATTACT  
AGATTTAAGG TGAATATATA GATTAAAGGT GAACTAGGGA AAATACCTAT TTAAGACCTA TTGTTAATGA

780 790 800 810 820 830 840  
ACCAGCAGTA TATCCTACTA TCAGCGCACT GCACACCAAA CTACCCTCAC CCAGTAGTTA CAAACGCATA  
TGGTCGTCAT ATAGGATGAT AGTCGCGTGA CGTGTGGTTT GATGGGAGTG GGTCAATCAAT GTTTGCGTAT

**Fig. 2A**

850 860 870 880 890 900 910  
 TTTTGCCGTT AGTTAATTAT TATCCGGTAA AGAAGGTAAA GAAGATTGGT AGTAATCCAA AATTTTCCCA  
 AAAACGGCAA TCAATTAATA ATAGGCCATT TCTTCCATTT CTTCTAACCA TCATTAGGTT TTAAAAGGGT

920 930 940 950 960 970 980  
 ACCCCAACCT CGGAACAAAA ACCGCGTAGT ATTTGTCGTA ACCAGGAGCA TCCGAGTCAT TAATTTACAC  
 TGGGGTTGGA GCCTTGTTTT TGGCGCATCA TAAACAGCAT TGGTCCTCGT AGGCTCAGTA ATTAATGTG

>Transcription\_start\_site

>CAAG\_site

990 1000 1010 1020 1030 1040 1050  
 CCAAACACAA AAAATTAGCA GCACGCAGCC GCCTTCCCAA TCCTCTCCTC TCTCCTCTCC TCTTCTCCAA  
 GGT TGTGTT TTTTAATCGT CGTGCCTCGG CGGAAGGGTT AGGAGAGGAG AGAGGAGAGG AGAAGAGGTT

1060 1070 1080 1090 1100 1110 1120  
 GCGGCAATTC GCGCGAGGTT TTCTCCGATC AAACCCTCGA ATCCCCCCT CGCGAATCCA TCGGAGGGTA  
 CGCCGTAAAG CGCGCTCCAA AAGAGGCTAG TTTGGGAGCT TAGGGGGGGA GCGCTTAGGT AGCCTCCCAT

1130 1140 1150 1160 1170 1180  
 GCGCGCGAT CCGCGTCGGC GAGAGCGGAT TCCGATTCCG CG ATG GAG CGG GTG TTC TCC GTG  
 CGGGGCGCTA GCGCGAGCCG CTCTCGCCTA AGGCTAAGGC GC TAC CTC GCC CAC AAG AGG CAC  
 M E R V F S V>  
 \_a\_a\_EXON1\_a\_a\_>

1190 1200 1210 1220 1230 1240  
 GAG GAG ATC TCC GAC CCA TTC TGG GTC CCG CCT CCG CCG CCG CAG TCG GCG GCG GCG  
 CTC CTC TAG AGG CTG GGT AAG ACC CAG GGC GGA GGC GGC GGC GTC AGC CGC CGC CGC  
 E E I S D P F W V P P P P P Q S A A A>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

1250 1260 1270 1280 1290  
 GCC CAG CAG CAG GGC GGC GGC GGC GTG GCT TCG GGA GGT GGT GGT GGT GTA GCG GGG  
 CGG GTC GTC GTC CCG CCG CCG CCG CAC CGA AGC CCT CCA CCA CCA CCA CAT CGC CCC  
 A Q Q Q G G G G V A S G G G G G V A G>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

1300 1310 1320 1330 1340 1350  
 GGC GGC GGC GGC GGC AAC GCG ATG AAC CGG TGC CCG TCG GAG TGG TAC TTC CAG AAG  
 CCG CCG CCG CCG CCC TTG CGC TAC TTG GCC ACG GGC AGC CTC ACC ATG AAG GTC TTC  
 G G G G G N A M N R C P S E W Y F Q K>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

1360 1370 1380 1390 1400 1410  
 TTT CTG GAG GAG GCG GTG CTC GAT AGC CCC GTC CCG AAC CCT AGC CCG AGG GCC GAA  
 AAA GAC CTC CTC CGC CAC GAG CTA TCG GGG CAG GGC TTG GGA TCG GGC TCC CGG CTT  
 F L E E A V L D S P V P N P S P R A E>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

**Fig. 2B**

1420                    1430                    1440                    1450                    1460  
 GCG GGA GGG ATC AGG GGC GCA GGA GGG GTG GTG CCG GTC GAT GTT AAG CAG CCG CAG  
 CGC CCT CCC TAG TCC CCG CGT CCT CCC CAC CAC GGC CAG CTA CAA TTC GTC GGC GTC  
 A G G I R G A G G V V P V D V K Q P Q>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

1470                    1480                    1490                    1500                    1510  
 1520  
 CTC TCG GCG GCG GCG ACG ACG AGC GCG GTG GTG GAC CCC GTG GAG TAC AAC GCG ATG  
 GAG AGC CGC CGC CGC TGC TGC TCG CGC CAC CAC CTG GGG CAC CTC ATG TTG CGC TAC  
 L S A A A T T S A V V D P V E Y N A M>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

1530                    1540                    1550                    1560                    1570                    1580  
 CTG AAG CAG AAG CTG GAG AAG GAC CTC GCC GCG GTC GCC ATG TGG AGG GTACAGC  
 GAC TTC GTC TTC GAC CTC TTC CTG GAG CGG CGC CAG CGG TAC ACC TCC CATGTCG  
 L K Q K L E K D L A A V A M W R>  
 \_a\_a\_a\_a\_a\_a\_a\_a\_a\_a\_EXON1\_a\_a\_a\_a\_a\_a\_a\_a\_a\_a\_>

1590                    1600                    1610                    1620                    1630                    1640                    1650  
 CATTCTCCCC CCCTCTAGTA CTCGAGAGCT TACTGAGATC GGCAATGCTA GCTACTGTTT GCATCGAATG  
 GTAAGAGGGG GGGAGATCAT GAGCTCTCGA ATGACTCTAG CCGTTACGAT CGATGACAAA CGTAGCTTAC

1660                    1670                    1680                    1690                    1700                    1710                    1720  
 TTTATAGGTA TTTAGATCGG GCATTTCTAT AGACCAATGG CGTCCATGGT CTTGCAATGC GCTCTGTTGA  
 AAATATCCAT AAATCTAGCC CGTAAAGATA TCTGGTTACC GCAGGTACCA GAACGTTACG CGAGACAAC

1730                    1740                    1750                    1760                    1770                    1780                    1790  
 GTGTCGGTGG TTGGTTCGAC TCATAGTATG TAGGGTTGTG CGTATGTACA AACGGAAGCT TCATAGACCT  
 CACAGCCACC AACCAAGCTG AGTATCATAC ATCCCAACAC GCATACATGT TTGCCTTCGA AGTATCTGGA

1800                    1810                    1820                    1830                    1840                    1850                    1860  
 CGGTATTGAG ATTGCGATAT CGATGCAACC TGCGAATTGG CGATGTAATC AGTCATATTC TTAATAAAT  
 GCCATAACTC TAACGCTATA GCTACGTTGG ACGCTTAACC GCTACATTAG TCAGTATAAG AATGATTGTA

1870                    1880                    1890                    1900                    1910                    1920                    1930  
 GCGAGACAGT GGTTTGTGTT CAATTGCAAT ATTTTGTGAT GGGGCTGCTT AAAGTGTGCT TGCCTTTTTA  
 CGCTCTGTCA CCAAACAAAC GTTAACGTTA TAAAAACATA CCCCACGAA TTTGACAGTA ACGGAAAAAT

1940                    1950                    1960                    1970                    1980                    1990                    2000  
 GATTGGCAAT ATGTGACTTT ATGCAAGTAT TTGATTGGGC GGATCCAGGA AAAAAAGTT GGGGGGATTC  
 CTAACCGTTA TACACTGAAA TACGTTTATA AACTAACCCG CCTAGGTCCT TGTTTTTCAA CCCCCCTAAG

2010                    2020                    2030                    2040                    2050                    2060                    2070  
 AACATACCGA GTACACTGGC ATAAACACAT CATCTCAGTA TTAAACTATG CTAAATGCT ATTAAGAGAC  
 TTGTATGGCT CATGTGACCG TATTTGTGTA GTAGAGTCAT AATTTGATAC GATTTTACGA TAATTCTCTG

**Fig. 2C**

2080 2090 2100 2110 2120 2130 2140  
 CTTTAGCACC TCTTATCTTA TCAACCATGG TGAAAAAATT GAAGGGGGGA CTCAGGGGGG TATCCATGGG  
 GAAATCGTGG AGAATAGAAT AGTTGGTACC ACTTTTTTAA CTTCCCCCCT GAGTCCCCC ATAGGTACCC

2150 2160 2170 2180 2190 2200 2210  
 TCCGATGGGT GCAGGGGGGA CTGAGTCCCC CCTGCACCCA CGTTGAATCC GCCCTGGCAT GCGTATAAGC  
 AGGCTACCCA CGTCCCCCCT GACTCAGGGG GGACGTGGGT GCAACTTAGG CGGGACCGTA CGCATATTCG

2220 2230 2240 2250 2260 2270 2280  
 TGTCACAGCC ATTTCTAGGT GCTTGTGCTT AGTTGGGTGA TGTCAGCTTA ATTTGTCTTT TCTATGTCGT  
 ACAGTGTCGG TAAAGATCCA CGAACACGAA TCAACCCACT ACAGTCGAAT TAAACAGAAA AGATACAGCA

2290 2300 2310 2320 2330 2340 2350  
 CATCGATTTT CTAAGAAACG AAAAATAGCC TATTTATGTG CTCCAGAATT TGATGATCCC TGGCCCTTCA  
 GTAGCTAAAA GATTCTTTGC TTTTATCGG ATAAATACAC GAGGTCTTAA ACTACTAGGG ACCGGGAAGT

2360 2370 2380 2390 2400 2410 2420  
 TTTGCTGAAA TTAGCCTATT TGTTGGTTGC CCTTCAGTTT TTTCCCAGCT TATGTTGTTG CAATGTGTGG  
 AAACGACTTT AATCGGATAA ACAACCAACG GGAAGTCAAA AAAGGGTCGA ATACAACAAC GTTACACACC

2430 2440 2450 2460 2470 2480 2490  
 CTATGCCTCG TTTTGTGCCC TATAATTTAT TATTTGCAAT TCATTTTTGT ACATGACTTA AAATGACACT  
 GATACGGAGC AAAACACGGG ATATTAAATA ATAAACGTTA AGTAAAAACA TGTACTGAAT TTTACTGTGA

2500 2510 2520 2530 2540 2550 2560  
 AGAGCAACAT GCACTGATTG GTTATCCTAT AATCATTTAT GTAGTTCTGT TCATTTTATC ATGCTAGCTC  
 TCTCGTTGTA CGTGACTAAC CAATAGGATA TTAGTAAATA CATCAAGACA AGTAAATAG TACGATCGAG

2570 2580 2590 2600 2610 2620  
 ATGTCATTTT CATCTTCAG GCC TCT GGC ACA GTT CCA CCT GAG CGT CCT GGA GCT GGT TCA  
 TACAGTAAAA GTAGAAGTC CGG AGA CCG TGT CAA GGT GGA CTC GCA GGA CCT CGA CCA AGT  
 A S G T V P P E R P G A G S>  
 \_b\_b\_b\_b\_b\_b\_EXON2\_b\_b\_b\_b\_b\_b\_>

2630 2640 2650 2660 2670 2680  
 TCC TTG CTG AAT GCA GAT GTT TCA CAC ATA GGC GCT CCT AAT TCC ATC GGA GGTACTTA  
 AGG AAC GAC TTA CGT CTA CAA AGT GTG TAT CCG CGA GGA TTA AGG TAG CCT CCATGAAT  
 S L L N A D V S H I G A P N S I G>  
 \_b\_b\_b\_b\_b\_b\_EXON2\_b\_b\_b\_b\_b\_b\_>

2690 2700 2710 2720 2730 2740 2750  
 TCTTATCTGG TTACATTTTC AGATTGTTAT GAAACTACCC AAATATCCTG CACAATTGCA TGGGATTAAA  
 AGAATAGACC AATGTAAAAG TCTAACAAATA CTTTGATGGG TTTATAGGAC GTGTTAACGT ACCCTAATTT

**Fig. 2D**

2760	2770	2780	2790	2800	2810	2820
TTTTAGTTTC	TTTGAAATAG	AAGTAGAGTT	GTATTGCTGT	CACGTCATCA	AATAGTTC TG	AAGCTATGAA
AAAATCAAAG	AAACTTTATC	TTCATCTCAA	CATAACGACA	GTGCAGTAGT	TTATCAAGAC	TTCGATACTT
2830	2840	2850	2860	2870	2880	2890
TAAATAAGTT	CCGCATTTGT	TAGTGATTCT	TTGAACATTA	GAATTGTTAT	GCTTAAGTAG	ATAGGGTTAT
ATTTATTCAA	GGCGTAAACA	ATCACTAAGA	AACTTGTAAT	CTTAACAATA	CGAATTCATC	TATCCAATA
2900	2910	2920	2930	2940	2950	2960
GTTTGTTTGG	AGTTCCCTTA	AATCATTTCA	TTGCTGACTG	CCAGCTGGCA	GGAGCATTTG	TTGTTGCCTT
CAAACAAACC	TCAAGGGAAT	TTAGTAAAGT	AACGACTGAC	GGTCGACCGT	CCTCGTAAAC	AACAACGGAA
2970	2980	2990	3000	3010	3020	3030
GACCATGAAT	GAAGACCTTC	CTGTTCTGAG	TGCTCACAA G	AAAACATATT	TTGATTAATG	CACCTTGAAT
CTGGTACTTA	CTTCTGGAAG	GACAAGACTC	ACGAGTGTT C	TTTTGTATAA	AACTAATTAC	GTGGAACTTA
3040	3050	3060	3070	3080	3090	3100
CCTTAGGATC	TTGCAAAGAT	GGGCACTTAG	CTTTAGAATT	GAGTAGTACT	TAAATAGCTG	TTGTTATCAT
GGAATCCTAG	AACGTTTCTA	CCCGTGAATC	GAAATCTTAA	CTCATCATGA	ATTTATCGAC	AACAATAGTA
3110	3120	3130	3140	3150	3160	3170
GATTTGT CCT	G TAGTGAAAT	G TCGACAAAA	C AGGAATGCT	A CTTTTGACT	T CTGATATTT	C ATGCCTGGC
CTAAACAGGA	CATCACTTTA	CAGCTGTTTT	GTCCTTACGA	TGAAA ACTGA	AGACTATAAAA	G TACGGACCG
3180	3190	3200	3210	3220	3230	
TTTACTTATG	CTCTGTTTGG	AACATGGGCA	CATATCA GGC	AAT GCT ACT	CCA GTT CAA	AAC ATG
AAATGAATAC	GAGACAAACC	TTGTACCCGT	GTATAGT CCG	TTA CGA TGA	GGT CAA GTT	TTG TAC
				G N A T P V Q N M>		
				__c__c__c__EXON3	__c__c__c__>	
3240	3250	3260	3270	3280	3290	
CTA AGT GGC	CCA AGT GGG	GGA TCG GGC	TCA CAG TTG	GTA CAG AAT	GTT GAT GTC	CTT
GAT TCA CCG	GGT TCA CCC	CCT AGC CCG	AGT GTC AAC	CAT GTC TTA	CAA CTA CAG	GAA
L S G P S G G S G S Q L V Q N V D V L>						
__c__c__c__c__c__c__c__c__c__EXON3						
3300	3310	3320	3330	3340		
GTA AAG CAG	CCC ACC AGC	TCT TCA TCA	AGG GAG CAG	TCA GAT GAT	GAT GAC ATG	AAG
CAT TTC GTC	GGG TGG TCG	AGA AGT AGT	TCC CTC GTC	AGT CTA CTA	CTA CTG TAC	TTC
V K Q P T S S S S R E Q S D D D M K>						
__c__c__c__c__c__c__c__c__c__EXON3						
3350	3360	3370	3380	3390	3400	
GGA GAA GCT	GAG ACC ACT	GGA ACT GCA	AGA CCT GCT	GAT CAA AGA	TTA CAA CGA	
CCT CTT CGA	CTC TGG TGA	CCT TGA CGT	TCT GGA CGA	CTA GTT TCT	AAT GTT GCT	
G E A E T T G T A R P A D Q R L Q R>						
__c__c__c__c__c__c__c__c__c__EXON3						

**Fig. 2E**

3410 3420 3430 3440 3450 3460 3470  
AGGTGATC ATTCATTGCT TCCTTGTAAT ATAGATTCTG TACATAATTA ACCTACCTCG TCATGCATGC  
TCCACTAG TAAGTAACGA AGGAACATTA TATCTAAGAC ATGTATTAAT TGGATGGAGC AGTACGTACG

3480 3490 3500 3510 3520 3530 3540  
ATGTGTCCTA TTTTCACCTT AGCCCTTTCA GTTGGATTTC CACTTTTCATC CGGTAGCCTT TCAGTTTCCT  
TACACAGGAT AAAAGTGGAA TCGGGAAAGT CAACCTAAAG GTGAAAGTAG GCCATCGGAA AGTCAAAGGA

3550 3560 3570 3580 3590 3600 3610  
ATTGCATCGC ATATATGATC TTTTACCTAC CATATTAGTT CTCTGTGTGC CATACTCAGT GCTTAGTGTGTC  
TAACGTAGCG TATATACTAG AAAATGGATG GTATAATCAA GAGACACACG GTATGAGTCA CGAATCACAG

3620 3630 3640 3650 3660 3670 3680  
TCGAGCAAGA GAGGAATTTG TATGGCTATT ACACGTAGCA CTTTGCTCTC TACTTGTTTTA TTGACATAAG  
AGCTCGTTCT CTCCTTAAAC ATACCGATAA TGTGCATCGT GAAACGAGAG ATGAACAAAT AACTGTATTTC

3690 3700 3710 3720 3730 3740 3750  
CAATTTGGGA TGAATTAAAT CTGAGTTCAC ATCATATTCC TTATGTCACA AGTTTCTGAA ACCGATTGTA  
GTAAACCCT ACTTAATTTA GACTCAAGTG TAGTATAAGG AATACAGTGT TCAAAGACTT TGGCTAACAT

3760 3770 3780 3790 3800 3810 3820  
TCTAGTATCT GGTTGATGCA CCCCCATCTT GGATTTGCAA ATCAAAGTTA TACTCCCTAG AGAGCTTTAC  
AGATCATAGA CCAACTACGT GGGGGTAGAA CCTAAACGTT TAGTTTCAAT ATGAGGGATC TCTCGAAATG

3830 3840 3850 3860 3870 3880 3890  
CTTTCATAAA GCAATTACCC CAATAAACCA CGGATTTGAT AGCTATTGAC TATGATTACC AGAATTCATT  
GAAAGTATTT CGTTAATGGG GTTATTTGGT GCCTAAACTA TCGATAACTG ATACTAATGG TCTTAAGTAA

3900 3910 3920 3930 3940 3950 3960  
TGGCAGCTAT TTTCTCAATT TAAGTTTGGT ATTAGTCTCA GTTGGCTGTA AAATAATGTC ACGGTAGGGT  
ACCGTCGATA AAAGAGTTAA ATTCAAACCA TAATCAGAGT CAACCGACAT TTTATTACAG TGCCATCCCA

3970 3980 3990 4000 4010 4020 4030  
ACATGTATGT GCAGCATACA AGGTATGGGT GAGTTATGAT ATGGACAGTG TGTACACCCC ACATTTGCTC  
TGTACATACA CGTCGTATGT TCCATACCCA CTCAATACTA TACCTGTCAC ACATGTGGGG TGTAAACGAG

4040 4050 4060 4070 4080 4090 4100  
ACTAAATCA AAATATTCAA ACGTCACGTG ATGATATGGT GGATTGCATT ATACCTTGTA TTGTTTATTA  
TGATTTTAGT TTTATAAGTT TGCAGTGCAC TACTATACCA CCTAACGTAA TATGGAACAT AACAAATAAT

4110 4120 4130 4140 4150 4160 4170  
TGTTACTTGT GCTAGACAAT AATATAGGCT GTTCTTTTGG GTGATTTTGT ATGAAGATGT TGAGCAAGCA  
ACAATGAACA CGATCTGTTA TTATATCCGA CAAGAAAACC CACTAAAACA TACTTCTACA ACTCGTTCGT

4180 4190 4200 4210 4220 4230  
CTTCTCGATA TAATGCTAGT TTTGTTGACC TGTTC AGG AAG CAA TCC AAT CGG GAG TCA GCC  
GAAGAGCTAT ATTACGATCA AAACAACTGG ACAAGG TCC TTC GTT AGG TTA GCC CTC AGT CGG  
R K Q S N R E S A>  
\_d\_d\_d\_EXON4\_d\_d\_d\_>

**Fig. 2F**



4240                4250                4260                4270                4280                4290  
AGG CGC TCA AGA AGC AGA AAG GCA GCT CAC TTG AAT GAG CTG GAG GCA CAG GTGTGA  
TCC GCG AGT TCT TCG TCT TTC CGT CGA GTG AAC TTA CTC GAC CTC CGT GTC CACT  
R R S R S R K A A H L N E L E A Q>  
\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_EXON4\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_d\_\_\_>

4300                4310                4320                4330                4340                4350                4360  
TAGTTCACAT AGTTATTTC GATAAGACAT AAAATCCATA ATTACTGGCT ACTGACTTCA GTTATGGATT  
ATCAAGTGTA TCAATAAAAG CTATTCTGTA TTTTAGGATT TAATGACCGA TGACTGAAGT CAATACCATA

4370                4380                4390                4400                4410                4420  
TACTTGTTAC AG GTA TCG CAA TTA AGA GTC GAG AAC TCC TCG CTG TTA AGG CGT CTT GCT  
ATGAACAATG TC CAT AGC GTT AAT TCT CAG CTC TTG AGG AGC GAC AAT TCC GCA GAA CGA  
V S Q L R V E N S S L L R R L A>  
\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_EXON5\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_>

4430                4440                4450                4460                4470  
GAT GTT AAC CAG AAG TAC AAT GAT GCT GCT GTT GAC AAT AGA GTG CTA AAA GCA GAT  
CTA CAA TTG GTC TTC ATG TTA CTA CGA CGA CAA CTG TTA TCT CAC GAT TTT CGT CTA  
D V N Q K Y N D A A V D N R V L K A D>  
\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_EXON5\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_e\_\_\_>

4480                4490                4500                4510                4520                4530                4540  
GTT GAG ACC TTG AGA GCA AAG GT ATGCTATATA TGCCTTTTGC AATATGCATC CCATGGATTG  
CAA CTC TGG AAC TCT CGT TTC CA TACGATATAT ACGGAAAACG TTATACGTAG GGTACCTAAC  
V E T L R A K>  
\_\_\_e\_\_\_e\_\_\_EXON5\_\_\_e\_\_\_e\_\_\_>

4550                4560                4570                4580                4590                4600                4610  
CTACTTTGGC TTGTTTCAAA CTTTCAACGT GACTTGTTGTA CCCTGTTATT AGAAGAATAA TCCC GCCTAC  
GATGAAACCG AACAAAGTTT GAAAGTTGCA CTGAACACAT GGGACAATAA TCTTCTTATT AGGGCGGATG

4620                4630                4640                4650                4660                4670                4680  
CAT TACTCT TATAAATCAC CATTTGGCCA GTCCAAACAT GATTATTAAA TCAGGTCATC CTGAACATTG  
GTAATATGAG ATATTTAGTG GTAAACCGGT CAGGTTTGTA CTAATAATTT AGTCCAGTTA GACTTGTAAC

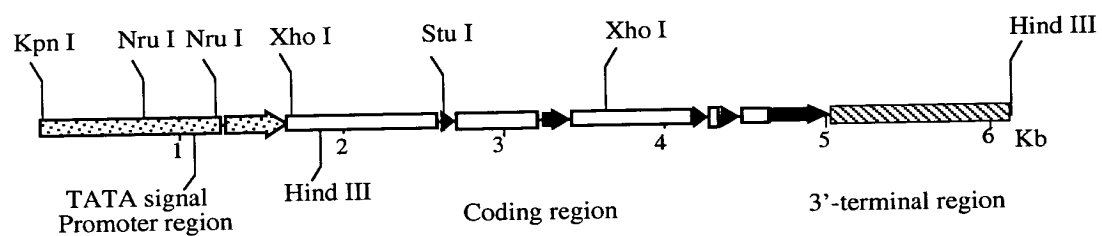
4690                4700                4710                4720                4730                4740  
AAATGTATCA AAAATTTCGCA GGTC AAG ATG GCA GAG GAC TCG GTG AAG CGG GTG ACA GGC  
TTTACATAGT TTTTAAGCGT CCAC TTC TAC CGT CTC CTG AGC CAC TTC GCC CAC TGT CCG  
K M A E D S V K R V T G>  
\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_EXON6\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_>

4750                4760                4770                4780                4790  
ATG AAC GCG TTG TTT CCC GCC GCT TCT GAT ATG TCA TCC CTC AGC ATG CCA TTC AAC  
TAC TTG CGC AAC AAA GGG CGG CGA AGA CTA TAC AGT AGG GAG TCG TAC GGT AAG TTG  
M N A L F P A A S D M S S L S M P F N>  
\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_EXON6\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_f\_\_\_>



5370	5380	5390	5400	5410	5420	5430
AATATTTTGA	TCGATGCTTC	CTCTTGCTCT	TTGCTCTTAA	GCAACCAAGC	ATAAAGATAT	CACTACCTTT
TTATAAACT	AGCTACGAAG	GAGAACAGAA	AACGAGAATT	CGTTGGTTCG	TATTTCTATA	GTGATGGAAA
5440	5450	5460	5470	5480	5490	5500
TGAGCTGTTC	ATTTGAAGTG	CAAAGCTAAG	CTCAATATCT	CAGGTGTTCA	TTTGAAGTTT	AAAGGTGAAC
ACTCGACAAG	TAAACTTCAC	GTTTCGATT	GAGTTATAGA	GTCCACAAGT	AAACTTCAA	TTTCCACTTG
5510	5520	5530	5540	5550	5560	5570
TGATAACAAA	CGTCAGGCTA	TGGTGAATGA	AGGGACGTGT	ACATCCCTAA	TACATGTCAT	TTTCATAATC
ACTATTGTTT	GCAGTCCGAT	ACCACCTACT	TCCCTGCACA	TGTAGGGATT	ATGTACAGTA	AAAGTATTAG
5580	5590	5600	5610	5620	5630	5640
AAATTAGTTG	ATGCATTTTC	ACCCAGAATC	CCATCACAGT	TCATCATACA	AGCAAGTGTA	GTTATTAATG
TTTAATCAAC	TACGTAAAAG	TGGGTCCTAG	GGTAGTGTC	AGTAGTATGT	TCGTTCACAT	CAATAATTAC
5650	5660	5670	5680	5690	5700	5710
GTAATTTTTT	CGTTTAGAGA	AAAAAAAAGG	AAGCCTTATA	TAAGATTAC	CGGTGGGGTG	TGAACAATAA
CATTTAAAAA	GCAAATCTCT	TTTTTTTTTC	TCGGAATAT	ATTCTAAGTG	GCCACCCAC	ACTTGTTATT
5720	5730	5740	5750	5760	5770	5780
TCAATGAATG	AGATCGCATC	CCGTAAGGGC	AGCCTAGCTA	GACAAAAATG	CATAAACTC	CGTATACCAA
AGTTACTTAC	TCTAGCGTAG	GGCATTCCCG	TCGGATCGAT	CTGTTTTTAC	GTATTTTGAG	GCATATGGTT
5790	5800	5810	5820	5830	5840	5850
CCACAACAAC	GCTTGCGCAC	GCGCTCAAAT	GGCAGCGACT	TCATCGCTTT	CGCGGGCAAG	AAACGAATCA
GGTGTTGTTG	CGAACGCGTG	CGCGAGTTTA	CCGTCGCTGA	AGTAGCGAAA	GCGCCCGTTC	TTTGCTTAGT
5860	5870	5880	5890	5900	5910	5920
AGTGATACAT	TGGCAGGGAA	CCACCAAAAG	AAGGCCATCC	AATCCAATCC	ACTCCAACGC	GGCATGGAAG
TCACTATGTA	ACCGTCCCTT	GGTGGTTTTT	TTCCGGTAGG	TTAGGTTAGG	TGAGGTTGCG	CCGTACCTTC
5930	5940	5950	5960	5970	5980	5990
ACAAGACAGA	TGATTACAG	CTATCTTCTG	CTTCTACAAG	TTTGATACTT	TGTACTGTCC	TTTCAGGGAA
TGTTCTGTCT	ACTAAGTGTC	GATAGAAGAC	GAAGATGTTT	AAACTATGAA	ACATGACAGG	AAAGTCCCTT
6000	6010	6020	6030	6040	6050	6060
AAAAGAGCAT	CAGATTAGTC	TGATCTCGGG	CGCGTTGAGT	TCTTGTTGGA	GATCTTGTTG	TGGAGTGGCA
TTTTCTCGTA	GTCTAATCAG	ACTAGAGCCC	GCGCAACTCA	AGAACACCCT	CTAGAACAA	ACCTCACCGT
6070	6080	6090	6100	6110	6120	6130
GGAGTGACGA	TCGGCTGCCC	CGTTTTCTTC	TACCGAAACA	TCGCCAGTAA	AGAAGCCAAA	AAGACAATAA
CCTCACTGCT	AGCCGACGGG	GCAAAAGAAG	ATGGCTTTGT	AGCGGTCATT	TCTTCGGTTT	TTCTGTTATT
6140	6150	6160	6170	6180	6190	6200
TACGGCAATG	GGGATCGCCC	ATCTGCATAA	AACATTGCAT	GACGGAAGTG	ATTAATACAA	GAATGACATG
ATGCCGTTAC	CCCTAGCGGG	TAGACGTATT	TTGTAACGTA	CTGCCTTGAC	TAATTATGTT	CTTACTGTAC
6210	6220					
TAAGCTGATA	ATTACGCGTG	CAAGCTT				
ATTCGACTAT	TAATGCGCAC	GTTCGAA				

**Fig. 2I**



**Fig. 3**

# Figure 4

BRAT

```

10      20      30      40      50      60      70      80      90
AAGCTTGCAT GCCTGCAGG AGGAGAGGG AGAGATGGT AGAGAGGAG AAGAAGAGGA GGGGTGACAA TGATATGTGG GCCATGTGGC

100     110     120     130     140     150     160     170     180
CCCCACCATT TTTTAAATCA TTCTTTTGT GAAACTGACA TGTGGGTCCC ATGAGAATTA TTATTTTTCG GATCGAATT GCCACGTAAGC

                                     >Reb_site1
190     200     210     220     230     240     250     260     270
GCTACGTCAA TGCTACGTCA GATGAAGACC GAGTCAAAT AGCCACGTAAG CGCCACGTCA CCAAAACC ACCATCCAAA CCGCCGAGGG

280     290     300     310     320     330     340     350     360
ACCTCATCTG CACTGGTGT GATAGTTGAG GGACCCGTTG TATCTGGTTT TTTCGATTGAA GGACGAAAAT CAAATTTGTT GACAAGTTAA

370     380     390     400     410     420     430     440     450
GGGACCTTAA ATGAACCTAT TCCATTTCAA AATATTCTGT GAGCCATATA TCCGTGGGCT TCCAATCCTC CTCAAAATTA AGGGCCCTTT

460     470     480     490     500     510     520     530     540
TAAATAAGAT AATGCGCTTC TTTCAGTCAC CCATAAAAGT ACAAAACTAC TACCAACAAG CAACATGCGC AGTTACACAC ATTTTCTGCA

550     560     570     580     590     600     610     620     630
CATTTCACC ACGTCACAA GAGCTAAGAG TTATCCCTAG GACAATCTCA TTAGTGTAGA TACATCCATT AATCTTTTAT CAGAGGCCAA

640     650     660     670     680     690     700     710     720
CGTAAAGCCG CTCTTTTATGA CAAAATAGG TGACACAAA GTGTTATCTG CCACATACAT AACTTCAGAA ATTACCCAAC ACCAAGAGAA

730     740     750     760     770     780     790     800     810
AAATAAAAAA AAATCTTTTT GCAAGCTCCA AATCTTGGAA ACCTTTTCA CTCTTTGCAG CATTGTACTC TTGCTCTTTT TCCAACCGAT

820     830     840     850     860     870     880     890     900
CCATGTCACC CTCAAGCTTC TACTTGATCT ACACGAAGCT CACCGTGCAC ACAACCATGG CCACAAAAAC CCTATAAAC CCCATCCGAT

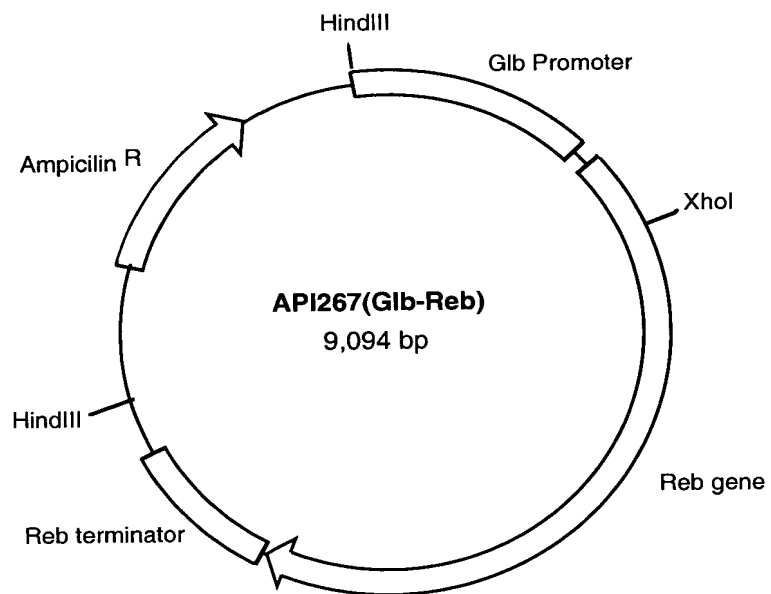
910     920     930     940     950     960     970     980     990
CGCCATCATC TCATCATCAG TTCATCACCA ACAAAACAAA GAGGAAAAA AACATATACA CTTCTAGTGA TTGTCTGATT GATCATCAAT

1000    1010    1020    1030    1040    1050    1060
CTAGAGGATC CCGGGGTGGT CAGTCCCTT ATG

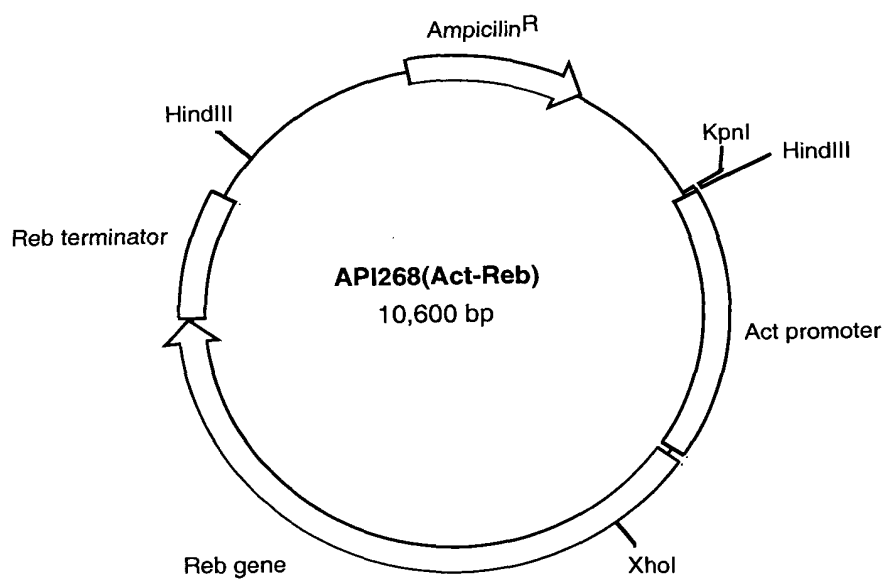
```

Gus Start site

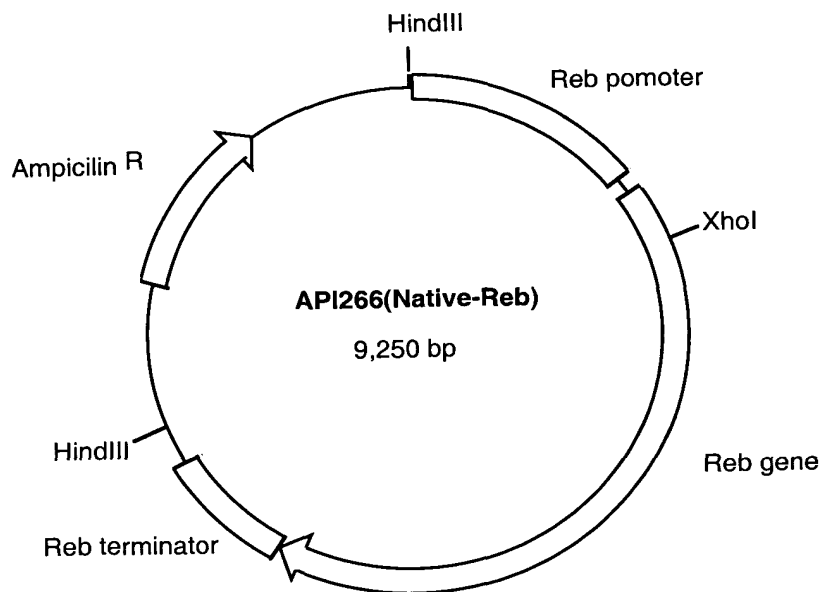
Fig. 4



**Fig. 5A**

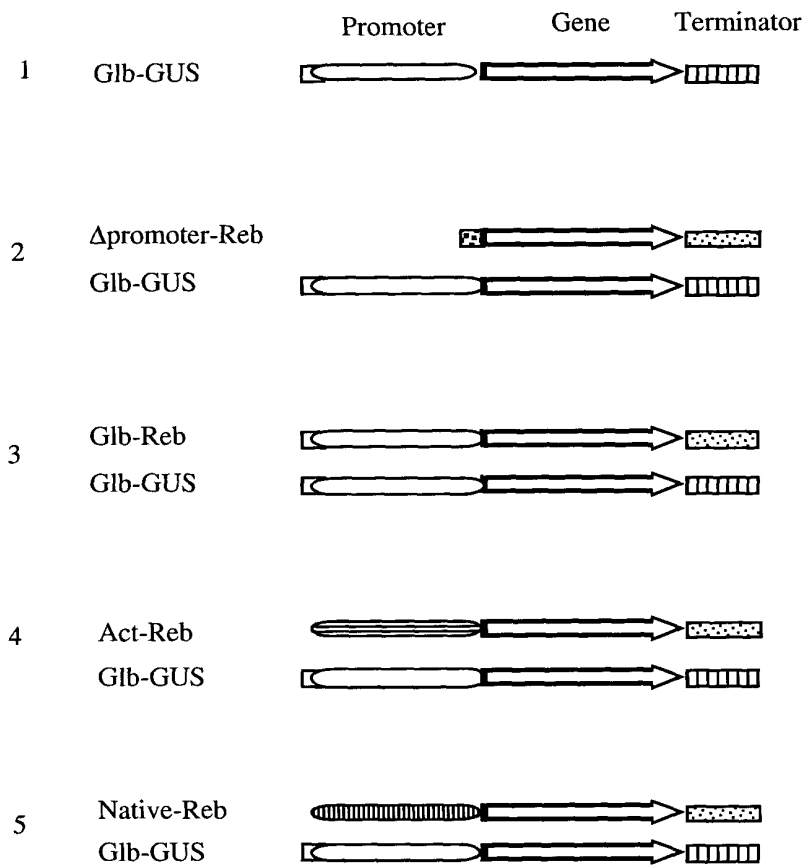


**Fig. 5B**

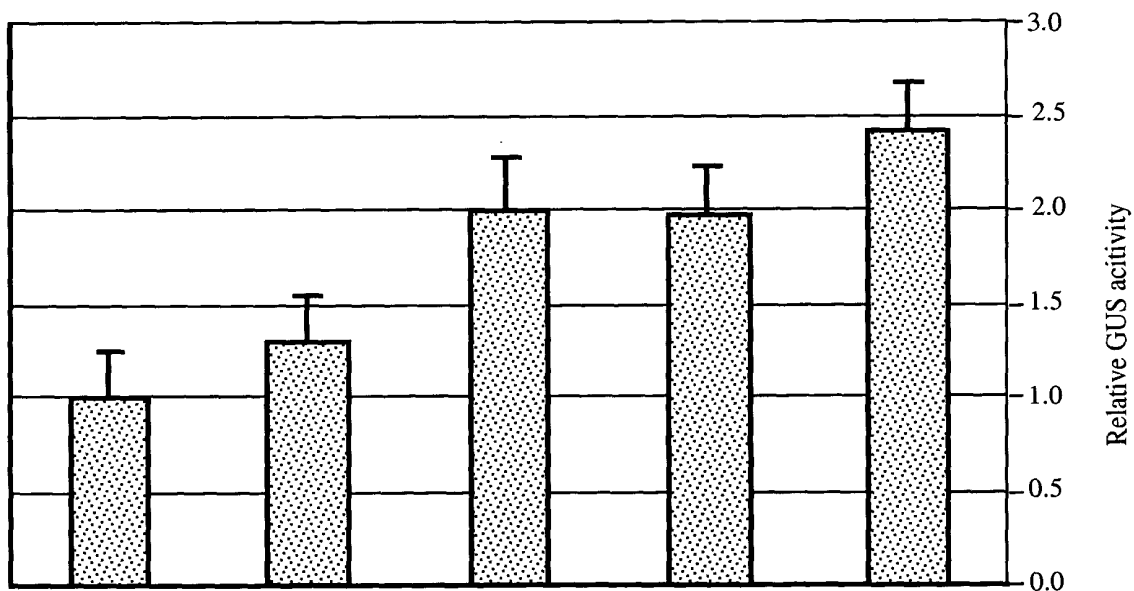


**Fig. 5C**

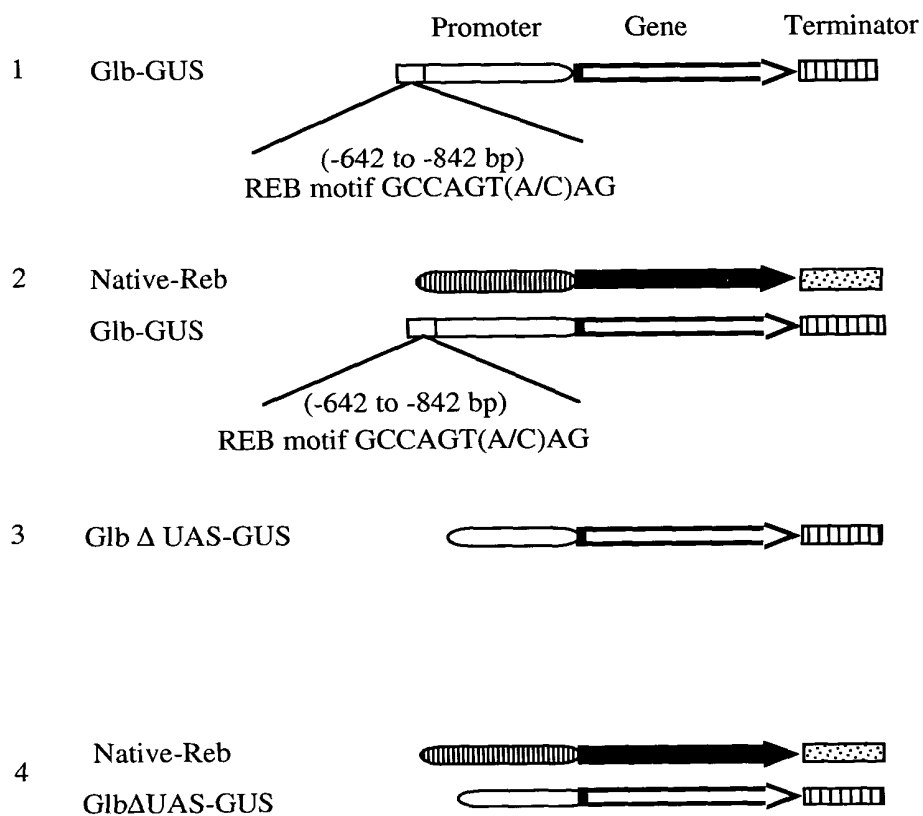




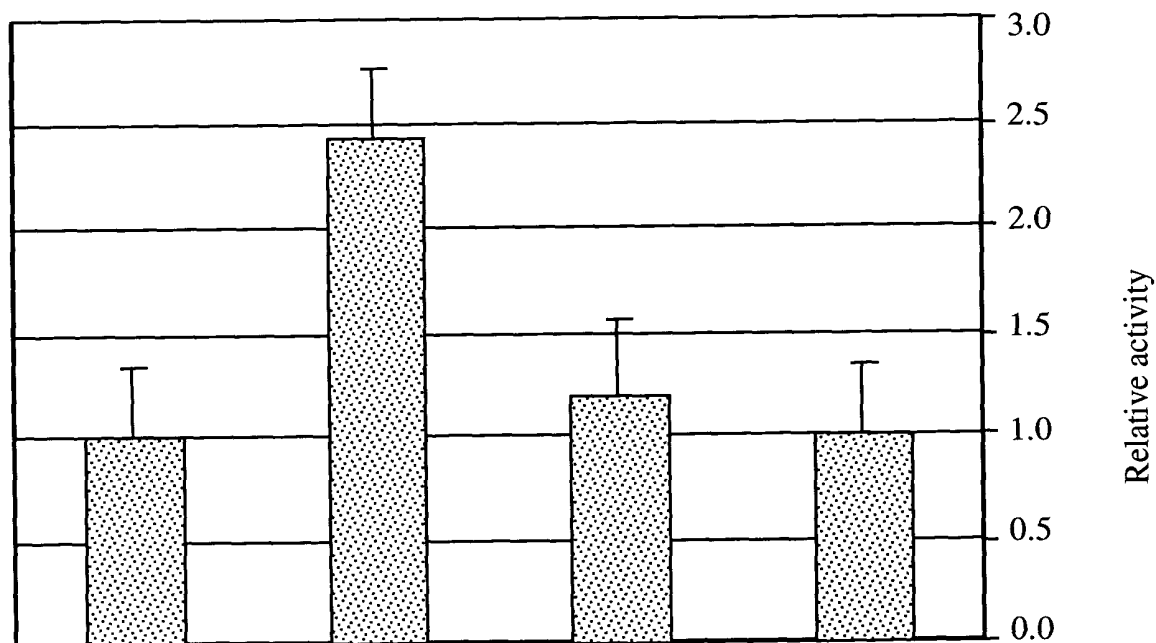
**Fig. 6A**



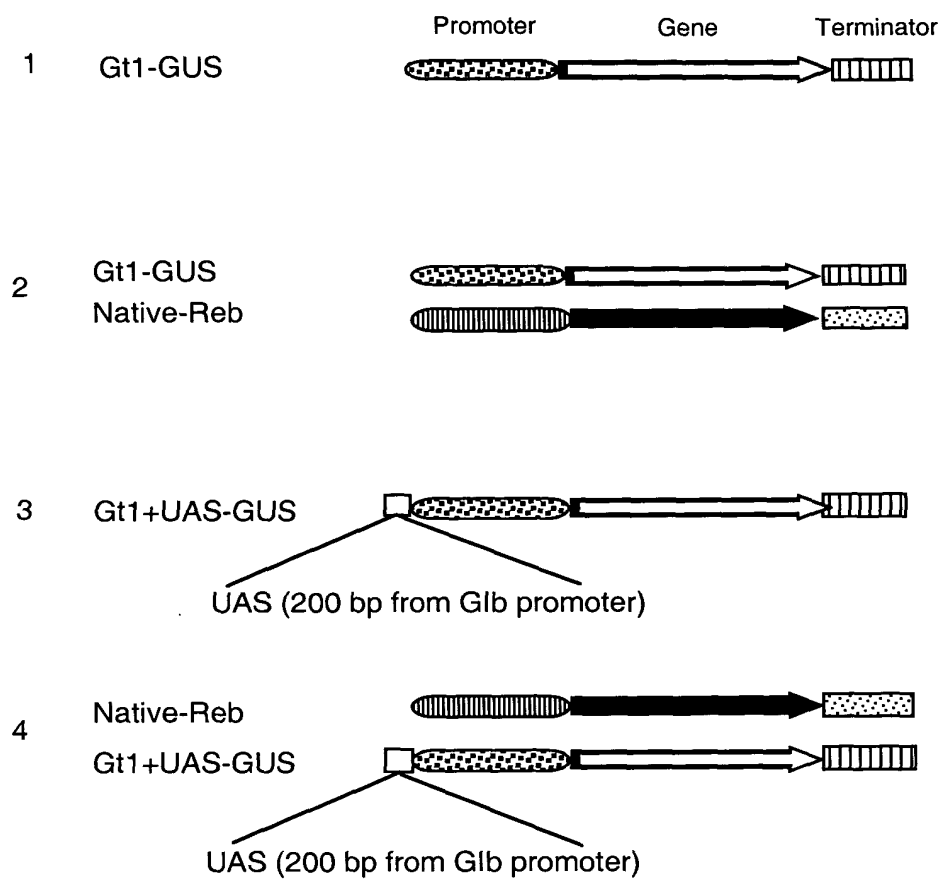
**Fig. 6B**



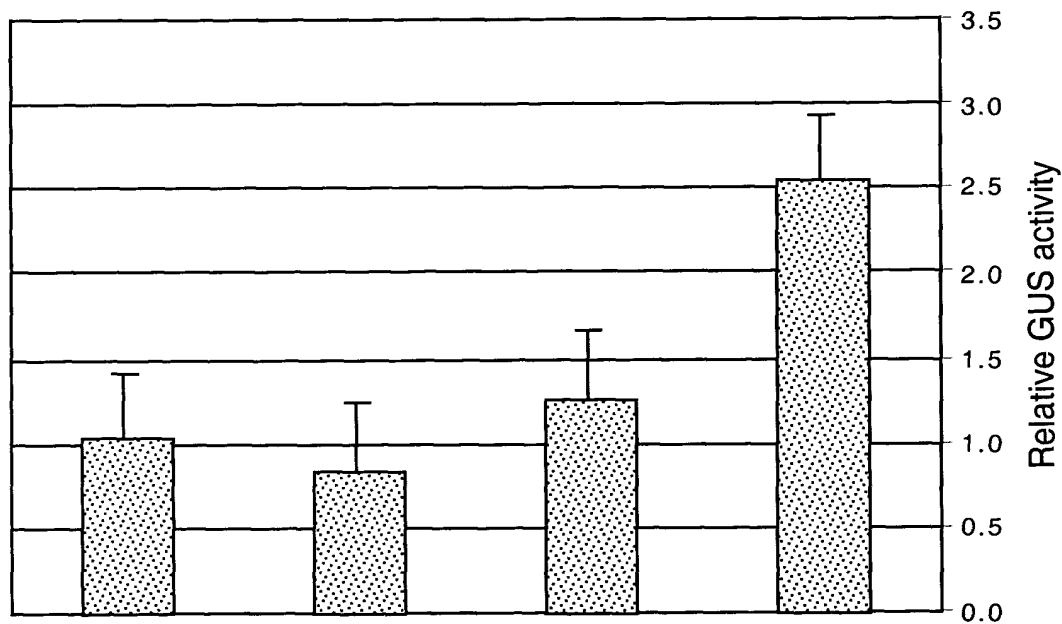
**Fig. 7A**



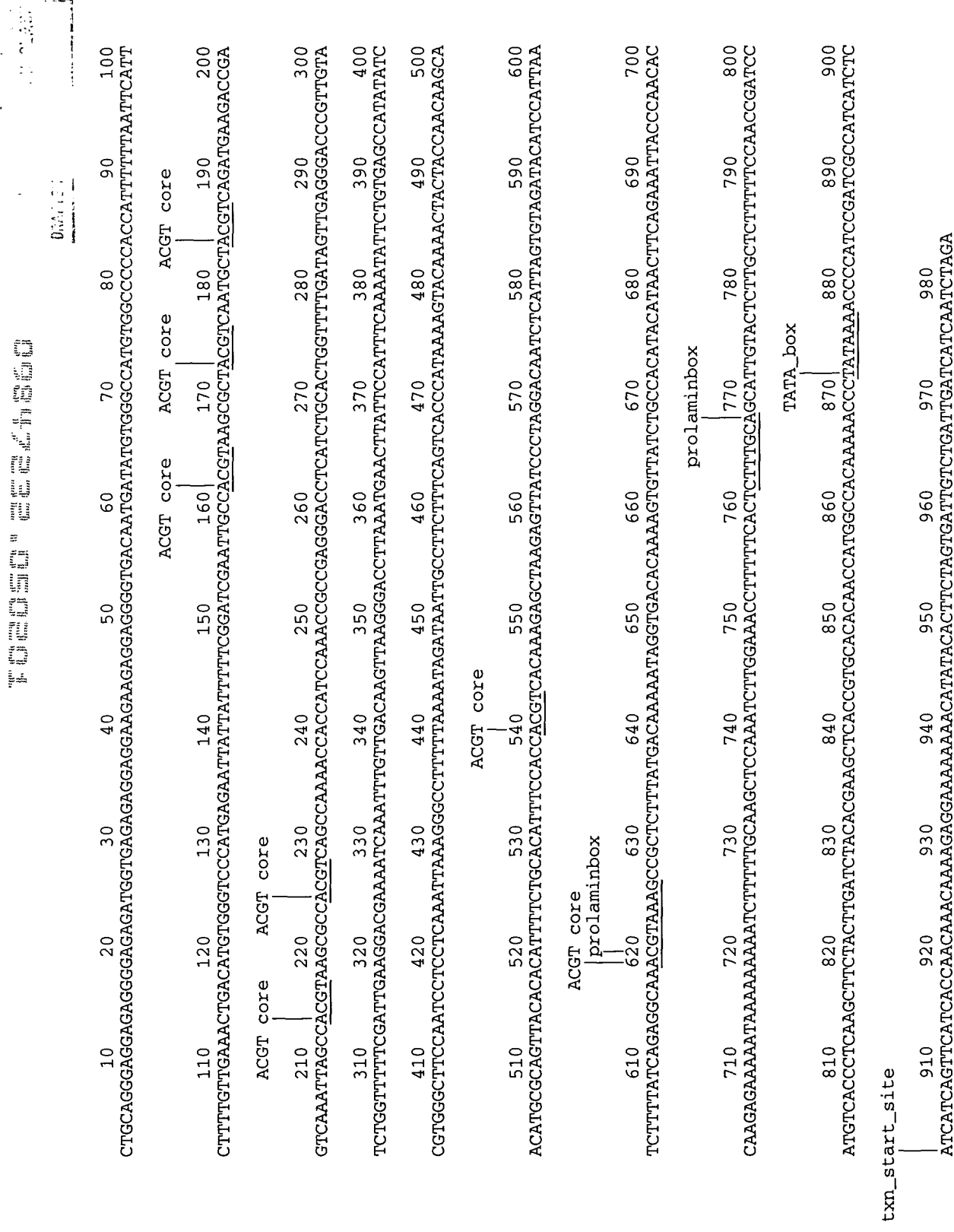
**Fig. 7B**



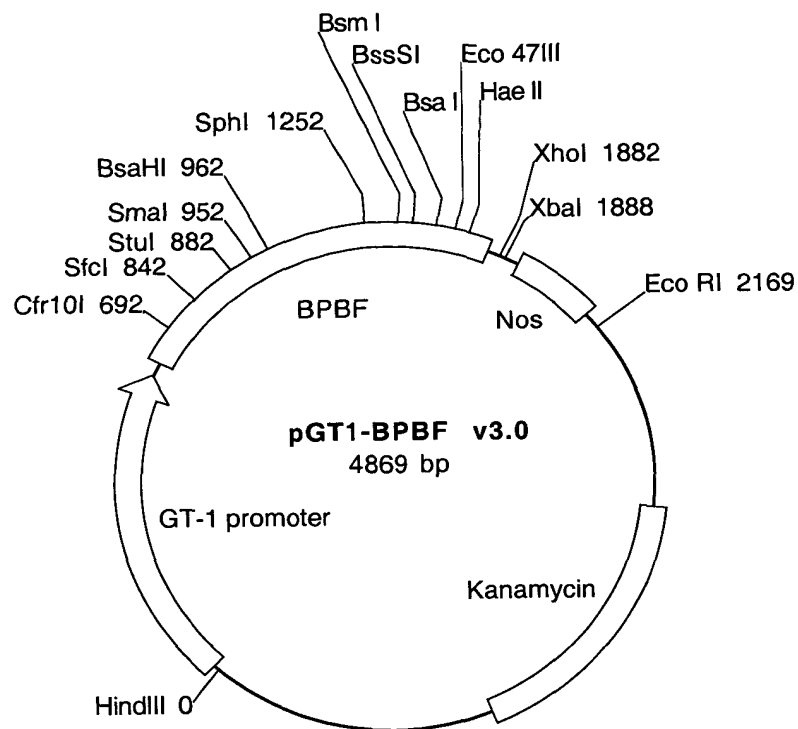
**Fig. 8A**



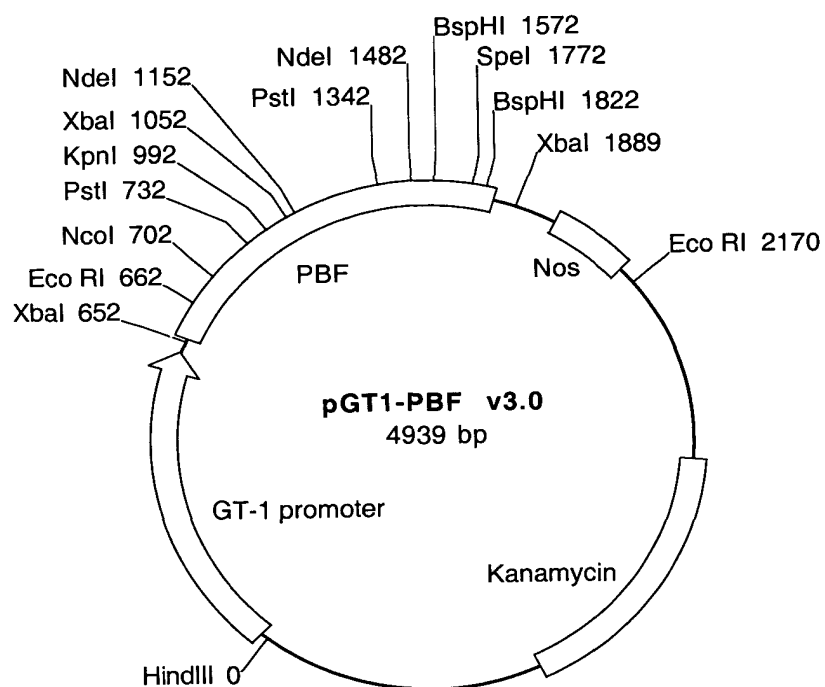
**Fig. 8B**



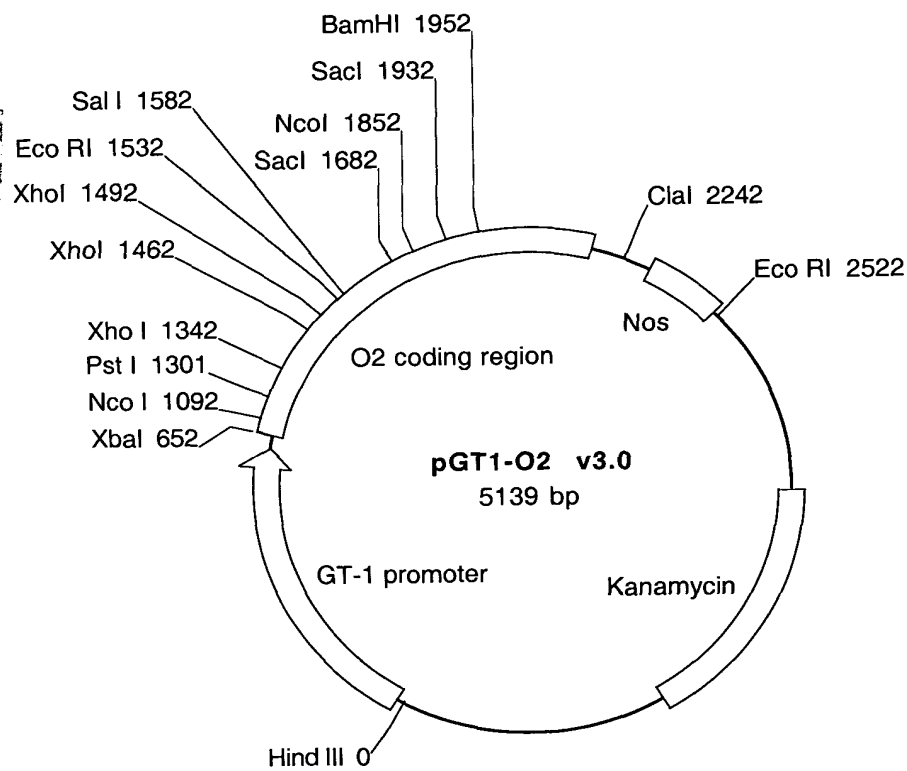




**Fig. 11A**

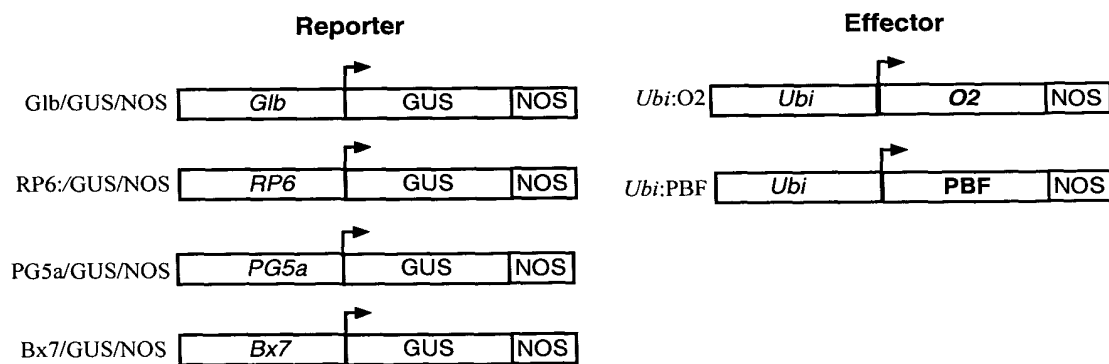


**Fig. 11B**

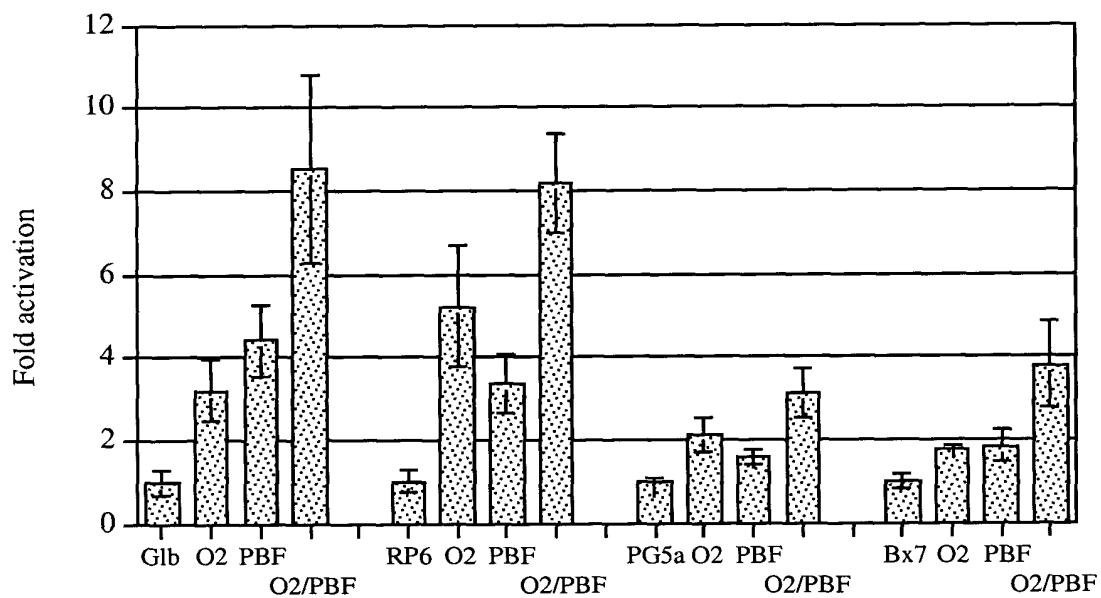


**Fig. 11C**

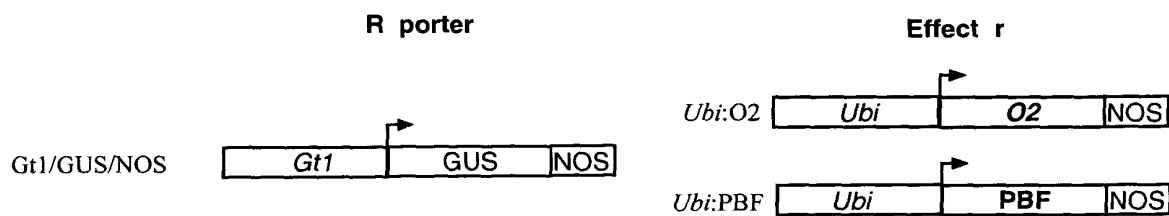




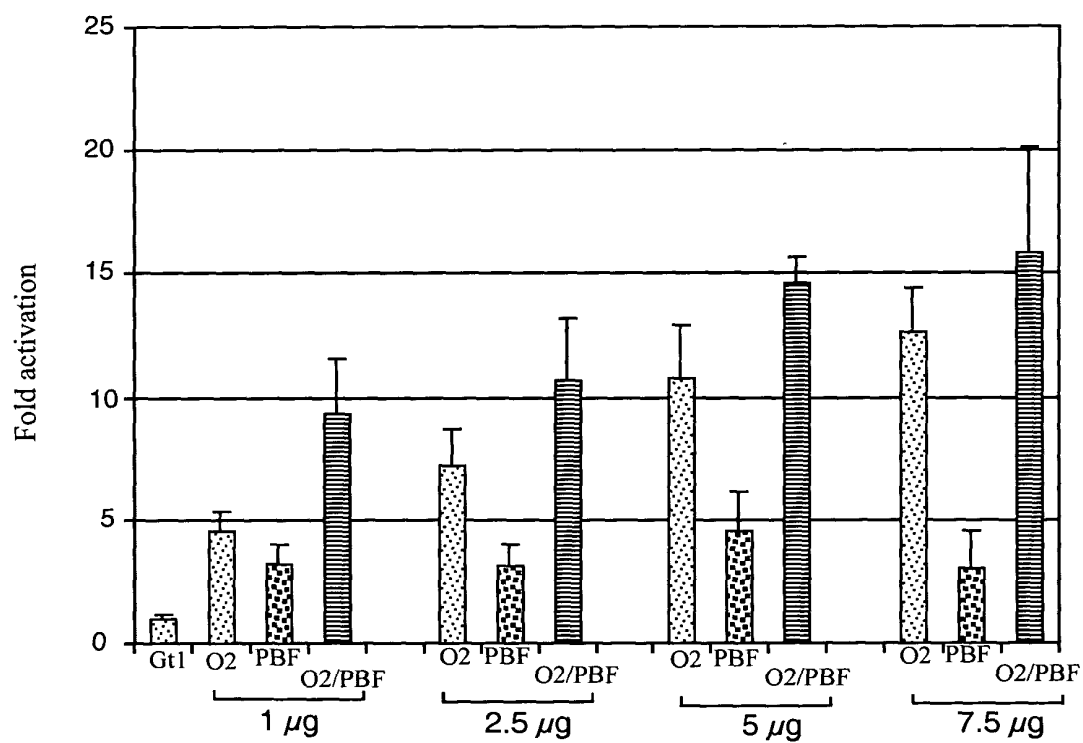
**Fig. 12A**



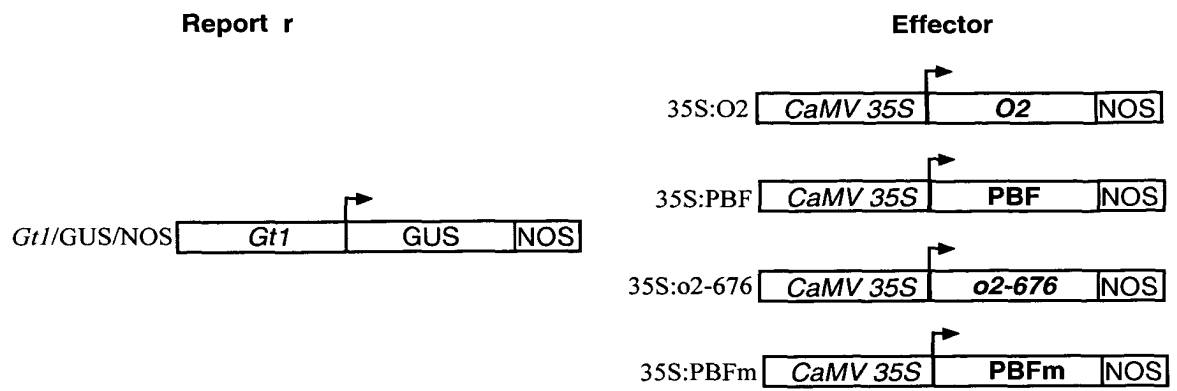
**Fig. 12B**



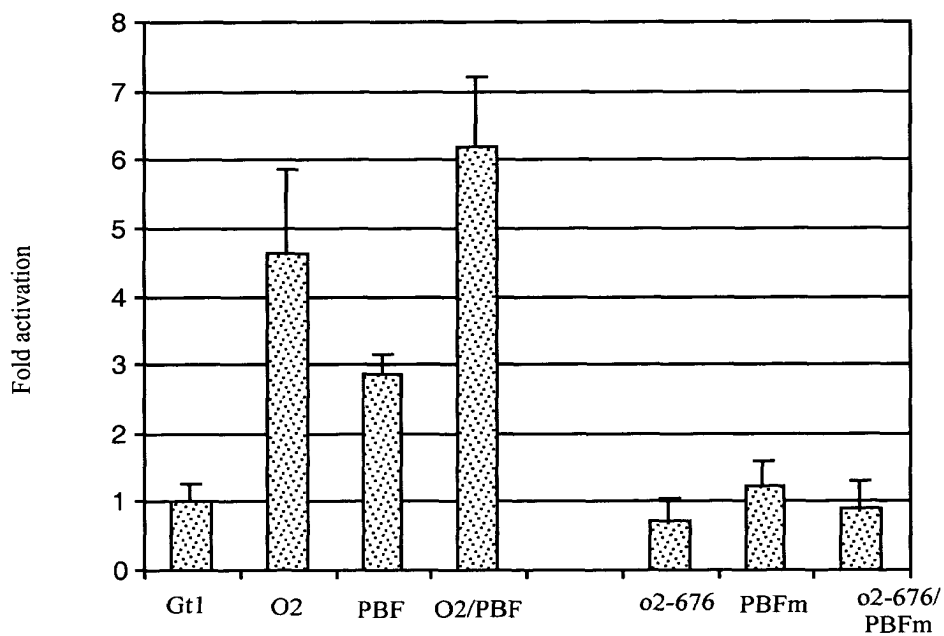
**Fig. 13A**



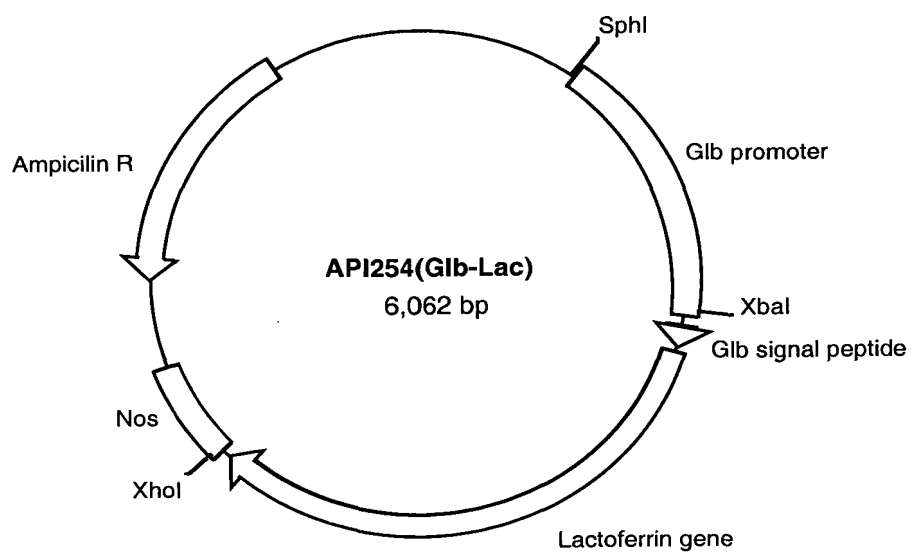
**Fig. 13B**



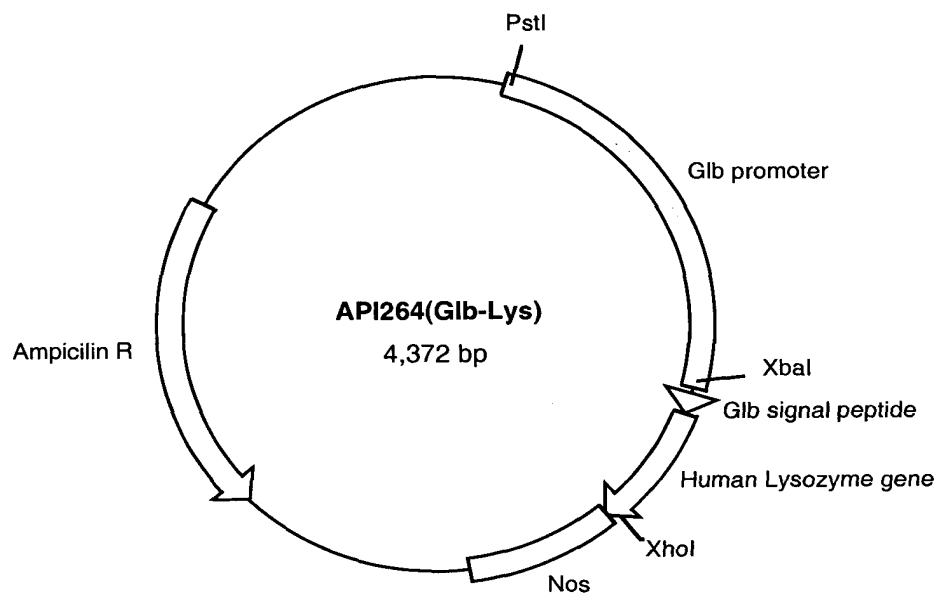
**Fig. 14A**



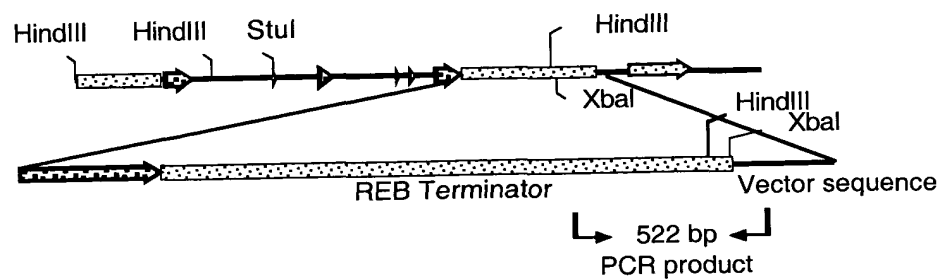
**Fig. 14B**



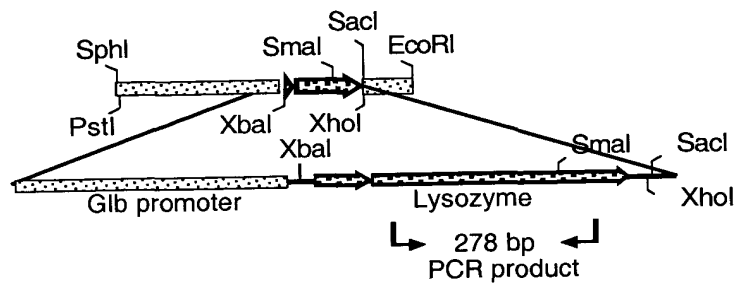
**Fig. 15A**



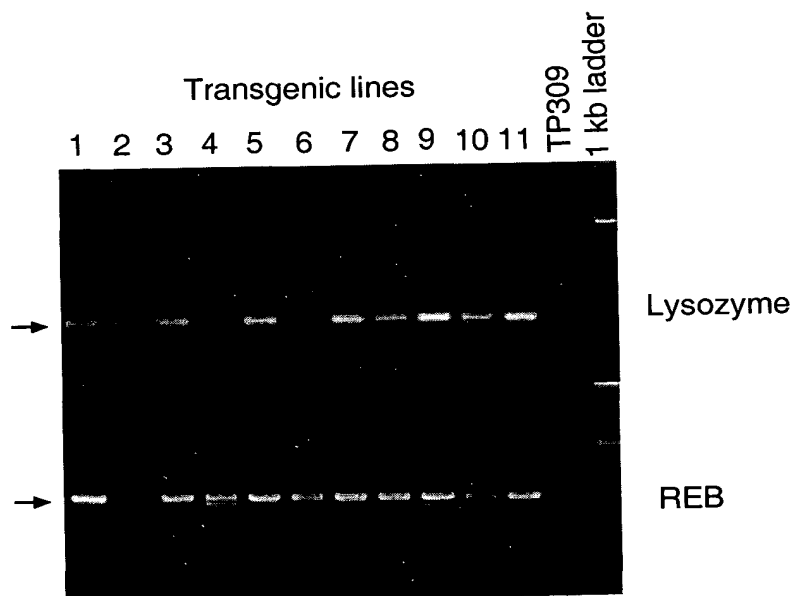
**Fig. 15B**



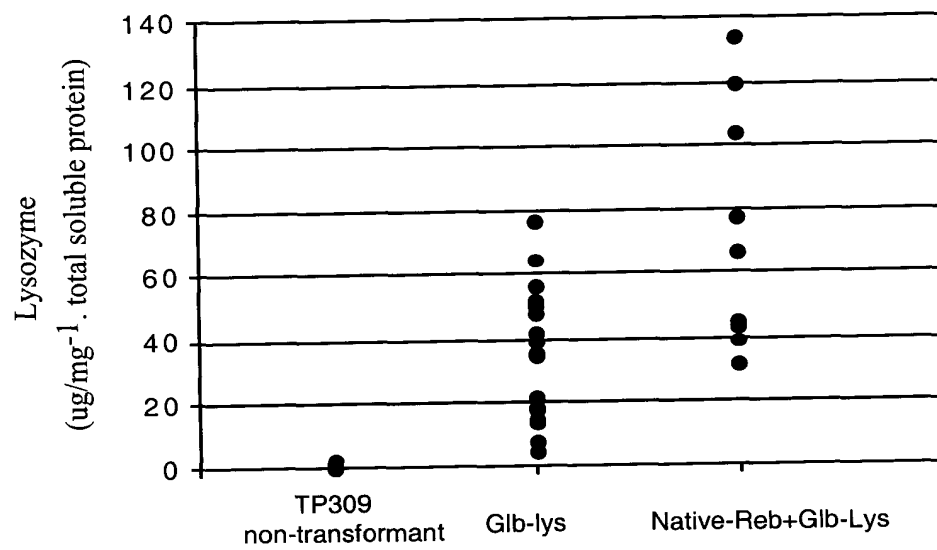
**Fig. 16A**



**Fig. 16B**



**Fig. 16C**



**Fig. 17**